EARLY REFLECTIONS IN MOBILE SOUND CONTROL ROOMS.

R. Walker Consultant.

1 INTRODUCTION.

Much has been written on the effects of early reflections and their control in audio monitoring environments over a period of about 30 years [e.g. 1-9]. Though there remain substantial disagreements amongst authors, this author at least is convinced that the effects of early reflections are real, potentially adverse and require proper consideration in the design of a room. This is especially true for the virtual images created in two-channel or multi-channel stereophony.

This paper considers the situation of high-quality monitoring environments in mobile sound control rooms. The severe constraints imposed on the acoustic design by the small space can result in significantly inferior sound quality at the primary listening position, especially with respect to spectral balance and virtual image localisation.

Measured results from a recent study of two particular vehicles are presented, together with interpretations of the underlying acoustic sound space and its artefacts. For comparison, calculated responses for the same geometry are also presented.

The investigation led to more fundamental insights into the effects of early reflections and has helped significantly to resolve some questions about the acoustic effects of early reflections.

2 THE AUDIBILITY OF EARLY REFLECTIONS.

2.1. General.

There is a large body of previous work on the audibility of relatively early 'echoes', essentially beginning with Haas in 1951 [10,11,12]. The majority of this work has been on isolated (or small numbers of) discrete echoes. More recent work has dealt with larger numbers of delayed sources, but the difficulties arising from the greatly increased number of degrees of freedom of such tests make interpretations difficult. Some more recent results [13] do have more relevance, though even in that case the numbers of simultaneous reflections that could simulated was limited. The relevance of these types of results to the stereophonic listening experience, in particular the creation of virtual or phantom centre images, in a room full of myriad reflections is less clear. There is also the fundamental problem of where to draw the boundary line between an effect detectable under stringent test conditions and one that materially alters the listening experience and the intent, i.e whether the listening is for pleasure or for production.

A recent, and very thorough, analysis of the factors involved in listening to reproduced sound in small spaces [14] essentially concluded that none of the hitherto assumed objective acoustic criteria for control rooms has much experimental justification, including the limits for early reflections. It is argued that, in the transition from large rooms (concert halls) for which they were originally derived to the entirely different conditions in small control rooms, so much has changed that their relevance has been lost. That author also makes some comments about the differences between listening for pleasure and objective programme production.

Despite that, this present author has remained convinced that the control of early reflections in production sound control rooms is important, even if there is uncertainty about the actual limits needed. In fact, the comments of Ref. 14 led this author to substantial introspection to try to resolve these differences. Ref. 14 makes quite clear that its author does not believe that early reflections cause adverse effects and even goes as far as to say that they may be beneficial. On the other hand, this author is convinced that severe adverse effects have been encountered in practice, observed by himself and complained of by production staff over many years.

The investigation described in this present paper has led to further insights and, perhaps, some reduction of this dichotomy. This is discussed further in Sections 3.4 and 6 below.

2.2. The distribution of early reflections in small and medium-sized rooms.

Partly in response to the comments in Ref. 14, a brief statistical study was carried out on the time and amplitude distributions of early reflections in typical control rooms. A simple computer programme was written to calculate the mean values and standard deviations for the relative time delays and amplitudes of the first-order reflections for a range of rectangular room sizes.

The programme applied simple design rules for the positions of one loudspeaker (left front) and the listener for a standard Rec. 775-1 layout [15]. The diameter of the loudspeaker-listener circle was adjusted to be as large as possible within the room constraints, up to a maximum of 2.5 m. The loudspeaker location was fixed at 1.0 m from the front wall and the height at 1.0 m. The listener height was fixed at 1.25 m.

The six first-order image locations were defined as :-

1 Front wall.
2 Righthand wall (ie. opposite side to the loudspeaker).
3 Rear wall.
4 Lefthand wall (ie. same side as the loudspeaker).
5 Ceiling.
6 Floor.

Fig. 1 shows the results obtained for smaller rooms. The range of room dimensions was from 5.0 to 7.0 m for plan and from 3.0 to 4.0 m for height. The number of rooms processed was 4851. The plot shows the mean image locations (large dots) and one standard deviation (ellipses) in time delay and amplitude relative to the direct sound. The plot also shows the nominal limits for early reflection control of 20 ms and 10 dB used by this author (and others).

As expected, the standard deviations for the floor (#6) and ceiling (#5) reflections were small. That would be expected from the almost fixed source and listener positions and the relatively minor effect of room height. The relationships between source, floor and listener were hardly dependent on room size. Similarly, the front wall reflection (#1) shows almost no variation because of the fixed loudspeaker spacing from that wall. The reflection from the wall nearest the loudspeaker (#4) also shows little variation, again because of the largely fixed geometry. All of these reflections have values for the mean ±1 standard deviation that lie within the 20 ms/10 dB design criterion.

The reflections from the wall on the far side from the loudspeaker (#2) and the rear wall (#3) had larger standard deviations – as expected because of the greater changes in reflection path length for different rooms sizes. Their mean values lay outside the criterion, even after subtracting one standard deviation.

That study did not include rooms typical of the size encountered in mobile facilities.

3 THE VEHICLES.

3.1. Overall.

Two vehicles were included in this study. In Vehicle A, the technical area consisted mainly of a single, long control room / recording area. It was used for many different programme genres, including complex continuity operations, sports, ‘pop’ and ‘classical’ music concerts. It had to include a wide range of facilities and space for several people. The control room / recording area had to be quite large and, because of the restricted width in vehicles, was inevitably rather long and narrow. The vehicle had been in use for a few years. Over that time there had been a substantial amount of criticism and a history of problems with the listening conditions.

Vehicle B was generally similar and used for a similar, fairly wide range of programme types. In that case, the ‘control room’ area was shorter and separated from the remainder of the vehicle by a partition. It also had physically different acoustic treatment, though the absorption performance and overall average reverberation times were essentially identical.

It was thought that the presence or absence of the partition was not having any significant effect on the stereophonic listening conditions. In both vehicles, reflections from the rear wall were considered to be outside the critical time/amplitude zone.

In both vehicles the main sound mixing desk was located near to end of the control room area and transversely across the width of the space. Because of the substantial size of modern mixing desks, in both vehicles the desks occupied all of the available width, touching the acoustic treatment on both side walls. The spaces behind the mixing desks were used for the stands for the main loudspeakers and any sub-woofers that might be in use.

3.2. The problems.

The sound mixing engineers using Vehicle A had reported a range of problems, of which only the difficulties with stereophonic (phantom) imaging are discussed in this paper.

Many criticisms had been made of the original choice of loudspeaker. Some regular users had developed their own preferences for alternatives. However, it was not clear which of the basic acoustic problems the changes in loudspeakers were intended to overcome. During the original investigation, no fundamental differences in the stereophonic image quality between different loudspeaker manufacturers or models were observed. The topic of loudspeaker selection is not considered further in this paper.

In Vehicle B, the sound mixing engineers were much more satisfied with the listening conditions. The vehicle was generally considered to be “adequate”, within the inevitable limitations of mobile facilities. The main reason for including the investigation in vehicle B was because of its reputation amongst the regular users for being acoustically much better than vehicle A.

The essential problems were to identify why these two apparently quite similar vehicles had such different listening conditions, to rectify Vehicle A and to try to ensure that similar problems did not recur in future, new vehicles

3.3. Acoustic design.

In Vehicle A, the control room area was finished, as far as possible uniformly, with the same type of acoustic treatment. That was a proprietary material consisting of dense or compressed mineral or glass wool in semi-rigid (self-supporting) sheets. The front surface consisted of a fairly hard fabric that was integrally bonded to the underlying material. It was not possible to investigate the structure of the treatment in more detail. There appeared to be no special provisions for bass absorption. Such as there was came only from the accidental absorption of the structure and fittings. The average reverberation time was uniform over most of the mid- and high frequency range. There was a substantial bass rise at 50Hz, but that was clearly not contributing to the stereophonic imaging problems. The average value of the reverberation time (RT) over most of the frequency range was about 140 ms. In terms of overall average reverberation, the vehicle was close to being anechoic. It is known that stereophonic listening in anechoic conditions works very badly [16].

Vehicle B had similar room geometry to vehicle A, although the control room area was not so long and was separated from the rest of the space by a sliding partition. The control room area was finished all over with quite soft, medium-pile carpet. The material underlying that was unknown but believed to be plywood. The ceiling consisted of what appeared to be proprietary perforated metal sheet material with a corrugated surface. The average mid-band RT was about 150 ms.

3.4. First impressions – main loudspeakers.

In Vehicle A, listening to both loudspeakers simultaneously with ‘commoned’ stereo feeds revealed a serious lack of clear, stable central images. By finding the optimum head position it was just about possible to obtain a kind of image, reasonably free from ‘phasiness’. Even then, it was more ‘in the head’ and not ‘in front’ as it should be. Even small head movements, of the order of 50mm, resulted in serious ‘phase’ effects and a complete breakdown of image resolution.

It was noted that these effects were definitely not evident with each of the loudspeakers driven individually. With more consideration (much later), it became evident that this precisely represented the differences in opinion described in Section 2.1 above. It was also the main justification for the preparation and presentation of this paper.

In Vehicle B, using the main loudspeakers gave a reasonably well-defined virtual centre image. That was obviously quite different to Vehicle A. Subjectively, the stereophonic image quality from the main loudspeakers was quite acceptable, with reasonably stable and consistent source image locations and a usefully wide ‘sweet spot’.

3.5. First impressions – near-field loudspeakers.

Both of the vehicles were equipped with near-field loudspeakers. As is normal, they were located on small shelves immediately above the outside rear corners of the mixing desks.

In both vehicles, the near-field loudspeakers presented a reasonably stable and well-defined virtual centre image. Listening in stereo, both loudspeaker systems produced a reasonably well-distributed sound stage with reasonably stable instrument and vocal locations, again within the inevitable limitations of mobile facilities.

4 GEOMETRIC MEASUREMENTS.

The physical geometries of the centre listening positions at the mixing desk were measured. Of course, it is not possible to be precise about the acoustic geometry – the acoustic position of a loudspeaker as a source is rather indeterminate. Other than on axis, the sound will be diffracted around the corners of the loudspeaker boxes and the path taken by a reflection via the wall or ceiling is difficult to estimate. Even the positions of the effective (acoustic) room boundaries are somewhat ill-defined because of the reactive impedance of the acoustic treatment and the diffraction effects as a result of the finite wavelength.

The geometric measurements were made to the voice coil of the mid-range unit for three-way loudspeakers or half way between the two drivers for two-way loudspeakers, to the measurement microphone diaphragm (representing the listing position) and to the physical front surfaces of the walls, floor and ceiling.

Appendix A.1 gives the significant primary dimensions for the two vehicles. As nearly as makes no practical difference, the rooms in the vicinity of the mixing desk were symmetrical about the front-back midline.

The main loudspeakers were very close to the side walls – in fact virtually touching them. Even so, that left a separation between left and right loudspeakers of only about 1.5/1.6 m. The loudspeakers were also only about 0.6/0.8 m from the ceiling. Limitations such as these are inevitable in the confines of a mobile facility.

5 ACOUSTIC MEASUREMENTS .

5.1. Equipment and general principles.

The acoustic measurements were carried using a B&K 2204 sound level meter with B&K 4134 (pressure/random incidence) microphone. An external, 'Purebits X4' USB sound card was used to capture the audio data. The sample frequency was 48kHz and all measurements were made with a 64 ksample record length. The data was processed by Sample Champion Pro® analysis software.

At the highest audio frequencies, even the ½" 4134 microphone becomes somewhat directional. When measuring direct sounds, the microphone can be orientated to provide at least an approximately level response. That is impractical for the measurement of multiple reflections in a room because they can arrive from any direction. For these measurements, the microphone was arranged so that the diaphragm was at an angle of about 45° to the floor, facing approximately the upper front corner of the room, centrally above the observation window to the studio. Responses at frequencies above about 10kHz would have been affected by the microphone directionality and should be interpreted with some care.

All measurements made were of the impulse response. Theoretically, that contains all of the information about the room response, within the limitations of the data acquisition system. The sampling parameters would theoretically restrict the measured frequency response to about 24 kHz, in practice more like 22 kHz with real anti-aliasing filters. The length of the time record (64 ksamples) limited the basic frequency resolution to about 1 Hz. The impulse responses were processed as required to derive time-domain and frequency-domain responses.

In previous work [5-9] it has been this author's practice to present measurements of Amplitude-Time-Frequency in the form of 3-dimensional 'waterfall' plots, an example of which (from other work) is shown in Fig. 2. In the present case, the timescales of the early reflection structures were too short to permit the individual reflections to be separated in that way. Therefore the details of the reflection structure had to be deduced (as best they could be) from frequency-domain responses derived from the time-windowed impulse response. That imposed a significant degree of frequency-domain 'smearing' due to the implicit convolution with a relatively wide frequency-domain response (the transform of the short time window). That effect is unavoidable in time-frequency measurements.

5.2. Outline of the investigation.

In this investigation, Vehicle A had been assessed and measured approximately 10 months before Vehicle B. The original plan had been to try to establish why Vehicle A suffered from poor stereophonic image quality. Based on those results and this author's extensive earlier work, it was assumed that that the excessive amplitudes of the early reflections in vehicle A were the main cause of the poor stereophonic listening conditions.

When the time came to carry out measurements in Vehicle B, it was anticipated that the soft and less regular front surface of the acoustic treatment in vehicle B would have provided improved image quality through lower amplitudes of high-frequency reflected sound.

Thus, this investigation can be divided into two consecutive parts – the first being an assumption (which turned out to be incorrect) and the second a revelation that the situation might not be as simple as that.

5.3. Early reflection results in Vehicle A.

Fig. 3 shows the overall amplitude response for the left loudspeaker at the central desk position, for the whole time record (64 ksamples = 1.33 s). It shows clear signs of a complex 'comb filter' response, obviously containing several discrete contributions. That is especially evident between 400 and 1200 Hz and between 2 kHz and 6 kHz.

Appendix A2, Table A.2.1 shows the calculated first-order early reflection arrival times (relative to the direct sound) and the predicted relative amplitudes (not including any surface attenuation) for the measured geometric arrangement. The reflections from the floor and the rear wall were obviously not significance.

Fig. 4 shows the response at the same position for just the first 1.5 ms (rectangular time window). Inevitably, the frequency resolution was severely limited by the short analysis time, but the contributions of at least two reflections are evident from the frequency spacing of the peaks and nulls. The dominant feature is a spacing of nulls at intervals of about 900 Hz, corresponding to a strong reflection at about 1.1 ms. There is some lesser evidence of a reflection at about 1.6 ms, indicated by the less well-defined peak/null structure with a spacing of about 600 Hz between about 3 kHz and 7 kHz. In the part of the frequency range where it was highest, the sum of the reflected sound energy (estimated from the average excess height of the ripples in the response) was about –4 dB relative to the direct sound[1]. That was a high level for early reflected sound energy (however, see Section 6.1).

Fig. 5 shows the response for the first 9 ms. The frequency resolution was rather better than in Fig. 4, but still limited to about 200 Hz. The contributions of nearly all of the significant early reflections were included in this range. Compared with the first 1.5 ms, the results show additional reflections at about 2, 3.3 and 5 ms. Again, in the frequency range where it was highest, the sum of the reflected sound energy was around –3 dB relative to the direct sound.

Associating measured reflection arrival times with calculated ones is not easy to because of the difficulties involved in making acoustically meaningful geometrical measurements (as above). However, it is likely that the earliest reflection, at about 1 ms, was from the left wall, the one at 1.5 ms from the ceiling, the one at 2 ms from the front wall and the one at 5 ms from the right hand wall. The floor reflection was obstructed by the mixing desk (as usual in control rooms) and the rear wall was too far away to be significant. In practice, reflection surfaces are more accurately identified on site by temporarily covering suspect surfaces with absorbing or reflecting material. Because of time restrictions, this could not be done at the time of the original investigation.

Comparing Figs. 3 and 5 shows that the basic 'comb filter' characteristic of the overall response was essentially fully developed by 9 ms after the arrival of the direct sound. That corresponds to path length differences, relative to the direct sound, of less than about 3 m, i.e. the first-order reflections from the immediately adjacent surfaces.

[1] It would be better to use the ratios of peak to null for calculating the relative levels of direct and reflected sound. However, the depth of the nulls was seriously under-represented because of the frequency domain smoothing inherent in the Fourier transform process used in the measurement equipment (see also Section 6.2). The peaks were not affected so much because of their relatively wider bandwidth.

6 MODELLING AND CALCULATIONS.

6.1. Discrete reflection amplitudes.

For some time, this author has been puzzled by the apparent contradiction presented by relatively large amplitude discrete reflections in heavily acoustically treated rooms. However, a simple calculation serves to show that this is actually not much of a contradiction.

If we take a space of the dimensions presented by the main control room area of this vehicle and its measured mid-band reverberation time, we can obtain the following :-

Room volume, *V*	= 38.65 m^3
Room surface area, *S*	= 80.6 m^2
Mean free path, *MFP* (= 4*V*/*S*)	= 1.92 m
Propagation time for 1 MFP	= 5.6 ms
Average number of boundary surface interactions in one reverberation time	= 25
Average attenuation at each interaction (=60dB / number of interactions)	= 2.4 dB

Thus we can see that, on average, even in this heavily treated room, we could expect each individual reflection to be attenuated by only 2.4 dB at the point of reflection. That does not include the attenuation due to the spreading loss nor any potential increases due to focussing or other effects from surfaces that are not quite flat. It is clear that quite high reflection amplitudes can be expected from room boundary surfaces, even in heavily treated rooms.

6.2. Calculated early reflection responses.

Theoretical amplitude responses were calculated, corresponding to the addition of a small number of discrete reflections to the direct sound. The results can be compared with the corresponding short-term responses actually measured.

Table 2 shows the values used for these calculations. In this case, the relative sound pressure level (Sound pressure ratio) does include the 2.4 dB average loss per reflection (Section 6.1).

Reflection surface	Front	Right	Left	Ceiling
Centre position (Figs 6 and 7)				
Relative time delay, ms	2.432	4.767	1.153	1.605
Sound pressure ratio	0.518	0.397	0.622	0.636
Left ear (Fig. 8)				
Relative time delay, ms	2.471	5.144	1.080	1.634
Sound pressure ratio	0.512	0.378	0.627	0.632
Right ear (Fig. 8)				
Relative time delay, ms	2.390	4.892	1.220	1.574
Sound pressure ratio	0.525	0.397	0.618	0.641

Table 2. Values of relative time delay and amplitude used for Figs. 6, 7 and 8

For these calculated responses, low-pass and high-pass filtering was applied to the delayed signals. That was intended to represent the attenuation of reflections at low frequencies because of wavelength (diffraction) effects and additional attenuation at high frequencies because almost all materials eventually become absorptive if the frequency is high enough. For these calculations, the weighting filter cut-off frequencies were 500 Hz (high pass) and 6 kHz (low-pass).

Fig. 6 shows the calculated response for a time limit of 1.5 ms. That corresponds with the parameters for the time-windowed measured response shown in Fig. 4 and includes just the ceiling and left wall reflections. Comparison of Figs. 4 and 6 shows that there are enough similarities in the general features of the responses at least to confirm the effect. In particular, it should also be noted that the calculated responses included individual reflection amplitudes as high as –4 dB. The calculated response includes apparently much larger irregularities than the measured one because the calculation was carried at closely-spaced discrete frequencies, whereas the measurement was made with an Fourier transform analyser. That implicitly included a substantial degree of frequency smoothing (convolution), corresponding to the 1.5 ms time window. That certainly prevented the measured nulls from being as sharp or deep as the calculated ones. The measured responses also included the loudspeaker response, which would add some more fine detail and certainly included a progressive lift from about 6 kHz, though the microphone directionality might have contributed to that also (Section 5.1).

It should be noted that the values of time delay and amplitude used for the calculations were the ones derived geometrically from the room dimensions and loudspeaker and listener positions and not those inferred from the measurements. That alone was likely to be responsible for some differences.

Figs. 7 shows the results of the same sort of calculation for a time limit of 9 ms. It corresponds with the parameters for the time-windowed measured response shown in Fig. 5 and included all four of the significant first-order reflections. It should be noted that, although the time window was 9 ms, the latest arrival time of the first-order reflections was about 5 ms. Thus, Figs. 5 and 7 more nearly represent the first 5 ms rather than the first 9 ms. Again, comparison of Figs. 5 and 7 shows that there are substantial similarities in the general shapes and some of the details of the responses. In this comparison, the effective frequency resolution of the measured response, Fig. 5, was better than for Fig. 4 because of the longer time window. The nulls in the measured response were not smoothed so much by the measurement system and consequently show as apparently somewhat deeper.

6.3. Effect of changes in receiver or listener positions.

It is instructive to compare calculated responses for slightly different receiver positions. That demonstrates the sort of response changes that would be encountered by a listener moving their head through small distances, as a studio engineer would during a session. It also indicates the sort of differences that might exist between the two ears of the listener, even without any head movement.

Fig. 8 shows calculated responses for all four of the significant first-order reflections at the two positions 100mm either side of the geometric centre position. That is just about the separation between the two ears of the listener. In reality, listening would be complicated by the two ears receiving signals from both loudspeakers, with the head providing some attenuation and phase shifting of the signals from the opposite side. For the purposes of this discussion (and because it is very hard to illustrate clearly because of the complexity of having four highly irregular graphs) that aspect of the listening has not been included here. (The purpose of this paper is to demonstrate the sort of artefacts that are encountered not a definitive prediction of any particular real listening experience.)

It should be noted that this author knows very well that the human hearing system does not directly perceive the irregular responses produced by normal steady-state frequency response measurements. Such measurements are, rightly, generally recognised as being of no great guide to the subjective sound quality. However, the results presented here are not steady-state responses (except for Fig. 3). They include only those features of the room response that arrive at the listener within a few milliseconds of the direct sound. Virtually all of the recognised literature concludes that the human hearing system does not discern as separate effects time-domain artefacts on that short a timescale. Rather, the early sound is integrated to provide an impression of the overall source 'quality'. Therefore these short-time measurements and calculated responses very likely do represent something like what the listener actually perceives.

At that stage in the investigation it was thought that the reason for the poor stereophonic images in Vehicle A had been identified. It is easy to imagine the severe disturbance of the perceived stereophonic virtual image (which, in the centre, is sensitive to amplitude variations of 1 dB or less) caused by the differences between the two ears represented by the two responses of Fig. 8. On that basis, it was thought quite understandable that the users were having difficulties in resolving central images and making sense of the virtual sound stage presented to them.

This situation is also not the same as loudspeakers themselves having irregularities in their anechoic response, as they undoubtedly all do. In that case, the irregularities remain reasonably constant and do not shift violently with small head movements and, at least for reasonable quality loudspeakers, are similar for the two channels. That was confirmed by the subjective impression with either one of the loudspeakers alone. That is also true of the effects of floor or ceiling reflections, as will be discussed later (Section 8).

7 INVESTIGATIONS IN VEHICLE B.

Long after the measurements and calculations described above had been completed (only three weeks before the deadline for submission of this paper), the opportunity arose to assess Vehicle B. It quickly became evident that the assumptions and interpretations described above had been, at best, incomplete.

Vehicle B, as a result of its rather more reflective ceiling treatment, had a more pronounced ceiling reflection than Vehicle A. Figs. 9 and 10 show the best attempts to measure the spectra of the opposite wall and ceiling reflections in the two vehicles[2]. Fig. 9 shows that in vehicle A, at frequencies above 2kHz, the reflection from the opposite side wall was generally very low, 12 – 15 dB below the direct sound and the ceiling reflection was generally around –10 dB. Fig. 10 shows that the reflections from the opposite side wall in Vehicle B were about the same as in Vehicle A but that the ceiling reflection was substantially higher, reaching within –5 dB of the level of the direct sound.

It cannot be the case that reflection amplitudes below about –20 dB have any significant effect in the presence of others at higher levels. Thus, it can be concluded that in Vehicle B the early reflection energy was dominated by the reflection from the ceiling. In contrast, all of the reflection amplitudes in Vehicle A were similar, individually quite low in amplitude and none were dominant.

8 NEAR-FIELD LOUDSPEAKER MEASUREMENTS.

Both vehicles were fitted with near-field loudspeakers in addition to the large main monitoring loudspeakers. In both vehicles, it was found that the near-field loudspeakers produced substantial acoustic reflections from the top surface of the mixing desk. However, the stereophonic image definition presented by the ‘near-field’ loudspeakers was much better than by the main loudspeakers. Indeed, it was entirely acceptable in both vehicles.

Using near-field monitors ought to result in an increased ratio of direct-to-reflected sound level. They are much nearer to the listener than the main loudspeakers and the reflection path lengths are not reduced in the same ratio, leading to more attenuation of the reflected sound because of the (relative) increase in distance. Also, the radiation angles corresponding to the reflection sound paths are more ‘off-axis’ than for more distant loudspeakers. Both of those effects ought to produce lower relative reflection amplitudes and, in turn, improved stereophonic images.

[2] There is a interesting contrast between the reflection amplitudes shown in Figs. 9 – 12 for individual reflections, which are generally at least –10 dB relative to the direct sound and Figs. 3 – 8, which include the summation of all of the early reflections at each frequency. In the summations, several reflections add their sound pressure levels at each frequency, each one dominating a different set of frequencies. The overall effect is to present a higher total level than any one reflection by itself. There is, therefore, no contradiction.

Fig. 11 shows the responses for the direct sound, the ceiling and the opposite side wall reflection for the left-hand near-field loudspeaker in Vehicle A. It can be seen that the relative ceiling and side wall reflection amplitudes were not very different in principle to those with the main loudspeakers.

However, Fig. 12 shows response of the reflection from the mixing desk top surface. In the high frequency region, it was around 5 - 7 dB below the direct sound, similar in amplitude to the ceiling reflection in Vehicle B (Fig. 10).

Vehicle B and all of the near-field loudspeaker arrangements had reasonable stereophonic image definition. Do we therefore conclude that ceiling or desktop reflections are necessary to the creation of a credible stereophonic virtual image in an otherwise highly absorbent environment? Certainly, Ref. 16 would support that.

Ceiling (or desktop or floor) reflections are different to those from other surfaces. Laterally, they are always in line with the direct sound. As such, they suffer no changes in relative response as the listener moves their head sideways. The irregularities remain reasonably constant and do not shift violently with small head movements and, at least for reasonable quality loudspeakers, are similar for the two channels. In that respect, they are no different to the direct sound from a real loudspeaker with its own response irregularities. It had already been established that a single real source (loudspeaker) in Vehicle A did not suffer from perceived response irregularities, though they were certainly present objectively. It is true that response changes would still occur as the listener moved vertically, but that is less common – and less easy to do without otherwise disturbing the listening situation.

9 DISCUSSION AND CONCLUSIONS.

It is reasonably clear that the differences between the two vehicles were in some way related to the early reflection distribution. On the basis of this work, we are faced with the inescapable conclusion that ceiling and desktop reflections (and floor reflections if present) assist in the presentation of stable virtual stereophonic images in an otherwise acoustically ‘dead’ environment. It is certainly true that the implication could be confirmed by the simple expedient of added a ceiling reflection to Vehicle A. At the time of writing there has been no opportunity to try that. Perhaps the critical condition is that on-axis reflections must dominate the early reflection pattern, whatever the general levels of early reflection amplitudes.

In this paper, and in many other references, the arrival directions, time delays and amplitudes of early reflections have been described (and indeed evaluated) using rather broad terms, usually assuming unidirectionality, smooth spectra and instantaneous arrival time. In reality, early reflections arrive from different directions, are uneven in spectral content and might even have frequency-dependant ‘arrival times’. It is perhaps unreasonable to assign broad limits of acceptability or effectiveness to artefacts that are so ill-defined. A wide field of further research is opened up, but the large number of potential degrees of freedom would probably make that difficult to carry out and then to apply in practice.

Though this paper has been a discussion in the context of a small control room in a vehicle, this author remains firmly of the opinion that the same sort of effects occur, perhaps to a lesser degree, in ordinary sizes of control room. In a reasonably designed, normal-sized control room, the earliest of the early reflections is likely to be close to the time limit at which the sound ceases to merge (≈ 5 ms). Their amplitude is also likely to be somewhat lower because of the additional spreading loss. However, in this author’s opinion it would not be appropriate to maintain that early reflections were not important to the perception of virtual stereophonic images and the preservation of the illusion of a consistent sound stage. It should be noted that this is quite different to the detection of adverse, or even beneficial, effects of early reflections on the perceived sound from a single, real source.

Historically, broadcasting sound engineers attempted to achieve well-defined and reasonably accurate stereophonic images. Many tests and experiments were carried out on the factors affecting the accuracy and resolution of the stereophonic sound stage. More recently, to some extent at least, the requirement for absolute accuracy has declined in importance – for many types of recording listeners do not have access to information about where the sources should appear to be located. In any case, many modern recordings are synthesised from solo tracks, entirely at the whim of the production team. However, for some classes of recording – for example orchestral music (where the listener might have an *a priori* expectation) or drama (where the intended location of individual sources is usually a crucial part of the production) – the absolute spatial accuracy is still important.

For nearly all kinds of production (with the possible exception of intentional *avant garde* compositions), whether or not the target audience needs to know the real source layout, it is still essential that the spatial consistency of each source is preserved. It is immensely confusing to the listener if an instrument or voice moves around the virtual sound stage with changes in pitch (unless it is intentional). The production team needs to have accurate and consistent monitoring to be able to achieve these objectives. The listener at home may be able to enjoy the production (at least to some extent) without the same sort of acoustic control, but that does not remove the requirement from the production process.

The conclusion reached in Ref. 14 that the effects of early reflections were either neutral or benign was based on the results from very many published papers on the subjective audibility of early reflections. However, without exception (as far as this author is aware), all of that historical work has been carried out using either loudspeakers or other primary sources that were single and real. The acoustic reflection patterns used for those tests were usually generated by secondary loudspeakers, with time delay, amplitude modification and possibly frequency-domain filtering and geometrical offsets. That is not at all the same as trying to produce a uniform, consistent and virtual stereophonic sound stage using an acoustic illusion based on two near-identical sources with small differences of amplitude and perhaps time delay.

Admittedly, at the extremes of the stereophonic sound stage, when the direct level differences might approach 10dB, the two are not so different, but the objectives are different. It could be argued that, at the extremes of the sound stage, the loudspeaker on the opposite side is deliberately trying to create an offset contribution rather like a room reflection might do (≈ –10dB) and if that didn't have any effect then the stereophonic sound stage effect would also not work out towards the extreme sides, which it clearly does. Generally, the time delays between the signals for the two loudspeakers in stereophony are equal to or close to zero. Because of that, the comb filter interference patterns are not generated to the same extent as for time-delayed early reflections but it is still true, and is indeed the basis of the stereophonic illusion, that the smaller signal from the opposite side loudspeaker is expected to affect the position of the virtual image. That is also one of the reasons why the two-channel stereophonic illusion does not work very well in large spaces. Once the loudspeaker spacing exceeds 4 – 4½ m, the path length differences become large enough to disrupt the illusion, presumably through the same sorts of mechanisms as early reflections. Even on a more compact scale, the same sort of phase addition (comb filter) effects as occur for early reflections may well be the main reason for the rather restricted 'sweet spot' in two-channel stereophony.

These measurements of very early and quite low amplitude reflections certainly explore the limits of what is possible with objective measurements. The presentation of these results has had to be selective in order to demonstrate, in a static form, the overall impression of what can be obtained by scanning the time windows and adjusting their lengths. The intention here has been to illustrate the general forms of the results rather than to produce definitive values.

The paper includes some significant and incompletely-resolved contradictions. In particular, why cumulative measurements over the first few milliseconds appear to indicate high levels of reflection and measurements of individual reflection amplitudes give much smaller values. Perhaps it is just the way the phases add to produce an apparently high level at some frequencies.

Some of the arguments presented in this paper are tentative. The author has not been able to carry out any other assessments to verify the results and is reluctant to overturn more than 20 years of work on the basis of one set of comparisons and with so little time to confirm the findings. The results have been presented as topics of discussion rather than as definitive conclusions.

In summary, whilst the original objective of this work was an investigation of problems with the monitoring conditions in one vehicle, the outcome has perhaps been a fuller understanding of the potential effects of early reflections on the perception of stereophonic virtual images. However, many new questions remain to be answered.

10 REFERENCES.

1. 1979, Walker, R. A preliminary investigation into the measurement of time and frequency response in listening rooms and control cubicles. BBC Research Department Report RD1979/9.
2. 1980, Davies, D. and Davies, C. The LEDE concept for the control of acoustic and psychoacoustic parameters in recording control rooms. J. Audio Eng. Soc., **28**, 9, pp. 585/595.
3. 1984, D'Antonio, P. and Konnert, J.H. The RFZ/RPG approach to control room monitoring. 76th Audio Engr. Soc. Convention, New York, October 1984, preprint #2157.
4. 1986, Muncy, N.A. Applying the reflection-free zone RFZ concept in control room design., dB The Sound Engineering Magazine, **20**, 4, pp. 35/39.
5. 1993, Walker, R., A new approach to the design of control room acoustics for stereophony. 94th AES Convention, Berlin, 1993, Preprint #3543.
6. 1994, Walker, R., Early reflections in studio control rooms: the results from the first installations. 96th AES Convention, Amsterdam, 1994, Preprint #3853.
7. 1994, Walker, R., Controlled image design: the management of stereophonic image quality. BBC Research and Development Department Report RD1995/4.
8. 1995, Walker, R., Controlled image design: results from the first installations. BBC Research and Development Department Report RD1995/5.
9. 1995, Walker, R., Controlled image design : the measurement of time-frequency responses. BBC Research Department Report RD1995/3.
10. 1951, Haas, H. The influence of a single echo on the audibility of speech. Acustica, 1, 2, p. 49.
11. 1954, Nickson, A.F.B., MUNCY, R.W. and Dubout, P. The acceptability of artificial echoes with reverberant speech and music. Acustica, **4**, pp. 447/450.
12. 1990, Fincham, L.R., Jones, A., Wilson, R.J. and Small, R.H. Free-field simulation of loudspeaker/ room acoustics. Proc. I.O.A., **12**, 8, Reproduced Sound 6, pp. 135/142.
13. 1995, Bech, S. Perception of Reproduced Sound: Audibility of Individual Reflections in a Complete Sound Field, II, 99th AES Convention, October 1995, Preprint Number 4093.
14. 2006, Toole, F.E. Loudspeakers and rooms for sound reproduction – a scientific review. JAES, **54**, 6, June 2006.
15. 1996, ITU BS.775-1. Multichannel stereophonic sound system with and without accompanying picture. International Telecommunication Union, Geneva, 1996.
16. 1974, Ratliff P.A. Properties of hearing related to quadraphonic reproduction. BBC Research Department Report RD1974/38.
17. 1984, D'Antonio, P. and Konnert, J.H., The reflection phase grating diffuser: design theory and application. JAES, 32, 4, April 1984, pp 228-236

11 ACKNOWLEDGEMENTS.

Fig. 3 was taken from earlier work, carried out while the author was employed by the BBC at Research Department, Kingswood Warren. It is reproduced by kind permission of the BBC.

APPENDICES

A.1 Main listening position and room dimensions.

Taking the front left lower corner of the room as the origin, the positions of the various features (in the order length, width, height) were :-

	Length	Width	Height
Room	8.18,	2.25,	2.1
Left main loudspeaker	0.48,	0.35,	1.53
Left near-field loudspeaker	1.20	0.28	1.28
Operator position at the mixing desk	2.05,	1.125,	1.25

Table A.1.1. Vehicle A.

	Length	Width	Height
Room	NA	2.25,	2.30
Left loudspeaker	0.60,	0.37,	1.55
Left near-field loudspeaker	1.20	0.47	1.21
Operator position at the mixing desk	2.23,	1.18,	1.25

Table A.1.2. Vehicle B.

A.2 Calculated early reflection delays and amplitudes.

Reflections:-		1	2	3	4	5	6
Surfaces -	**Direct**	**Front**	**Right**	**Rear**	**Left**	**Floor**	**Ceiling**
Total path length, m	1.79	2.62	3.43	13.88	2.18	3.22	2.34
Time difference, ms	0.00	2.43	4.77		1.15		1.60
Relative level, dB	0.00	-3.32	-5.64		-1.74		-2.33
Spl ratio (including average surface loss)	0.00	0.52	0.40		0.62		0.58

Table A.2.1. Calculated early reflection arrival times and levels in Vehicle A, left-hand main loudspeaker.

Surfaces -	**Front**	**Right**	**Left**	**Ceiling**	**Desk**
Path length distance, m	3.56	3.23	1.98	2.50	1.61
Time difference, ms	6.576	5.600	1.987	3.501	0.918
Relative level, dB	-8.73	-7.87	-3.66	-5.68	-1.88
Spl ratio (including average surface loss)	0.273	0.302	0.490	0.388	0.601

Table A.2.2. Calculated early reflection arrival times and levels – Vehicle B, left-hand near-field loudspeaker.

Surfaces -	**Front**	**Right**	**Left**	**Ceiling**	**Desk**
Path length distance, m	2.82	3.49	2.18	2.35	2.21
Time difference, ms	3.161	5.124	1.300	1.797	1.402
Relative level, dB	-4.22	-6.08	-1.99	-2.65	-2.13
Spl ratio (including average surface loss)	0.459	0.371	0.594	0.551	0.585

Table A.2.3. Calculated early reflection arrival times and levels – Vehicle B, left-hand main loudspeaker.

Surfaces -	Front	Right	Left	Ceiling	Desk
Path length distance, m	3.41	3.41	1.67	2.41	1.60
Time difference, ms	6.351	6.361	1.293	3.430	1.090
Relative level, dB	-8.86	-8.87	-2.68	-5.84	-2.31
Spl ratio (including average surface loss)	0.269	0.269	0.549	0.381	0.572

Table A.2.4. Calculated early reflection arrival times and levels – Vehicle A, left-hand near-field loudspeaker.

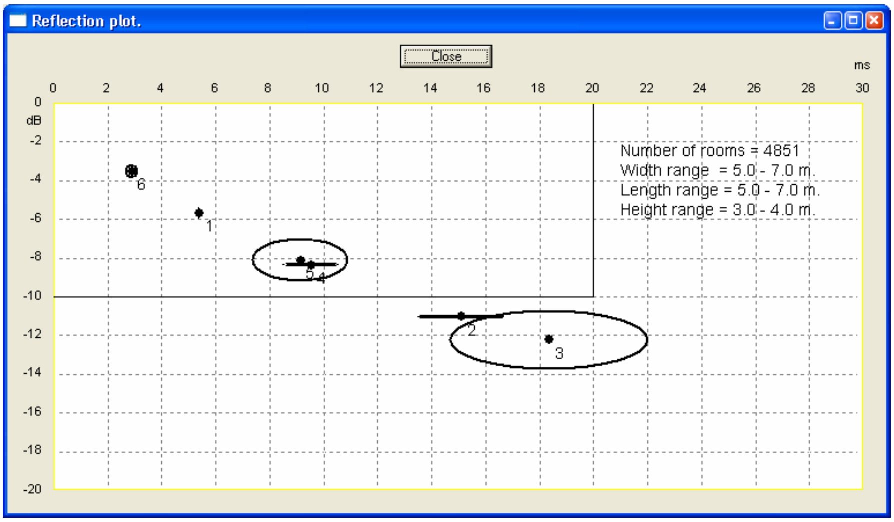

Fig. 1. Calculated reflection statistics for smaller rooms.

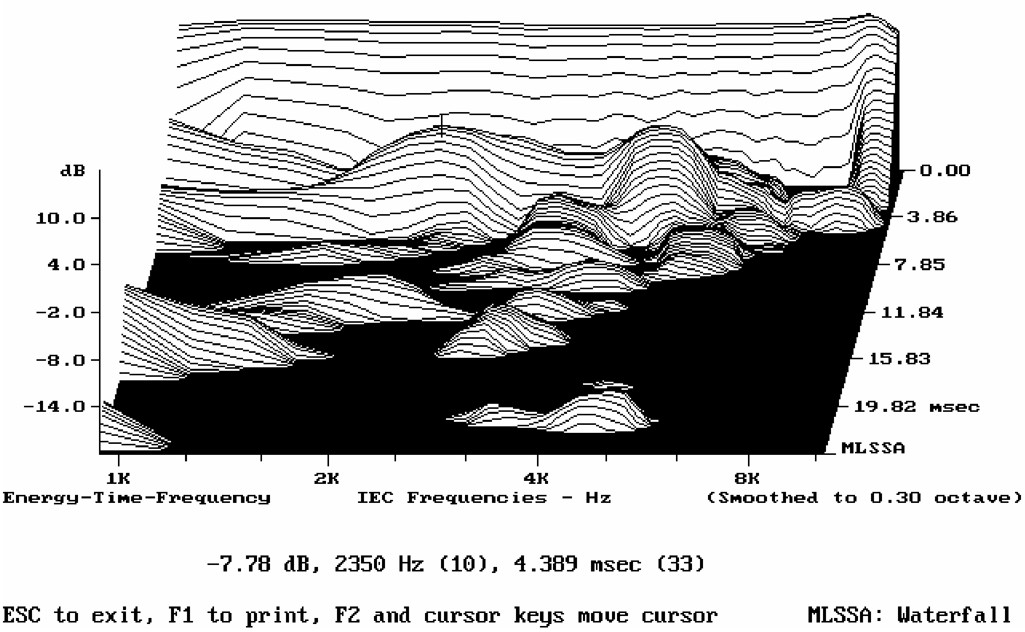

Fig. 2. 3-D plot from earlier work showing a pronounced early reflection at about 4.4 ms.

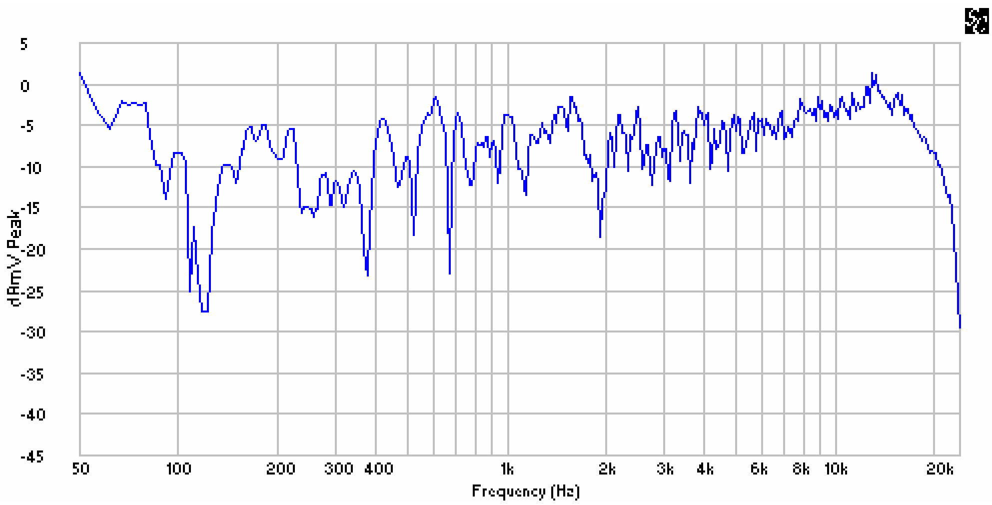

Fig. 3 Vehicle A - measured overall response for left-hand loudspeaker at desk operator position.

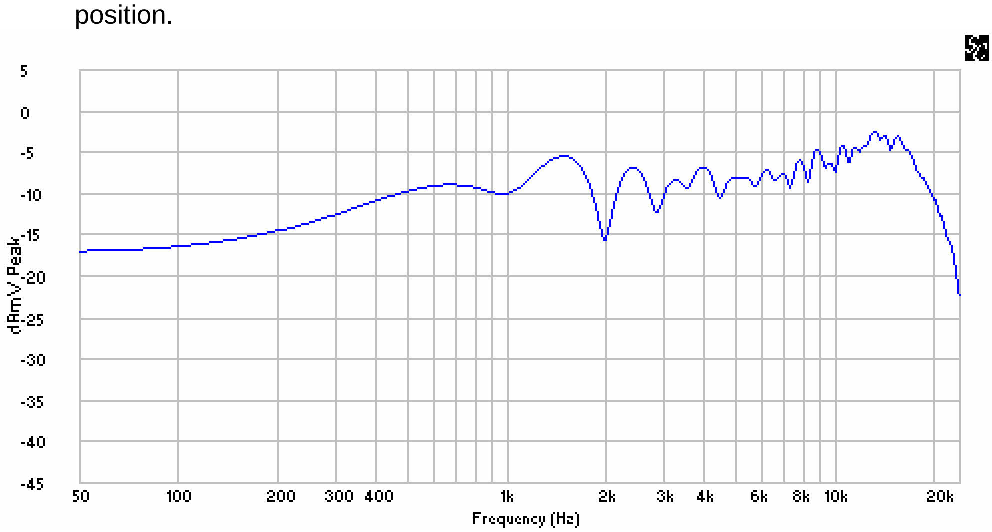

Fig. 4 Vehicle A - measured response for left-hand loudspeaker at desk operator position, first 1.5ms.

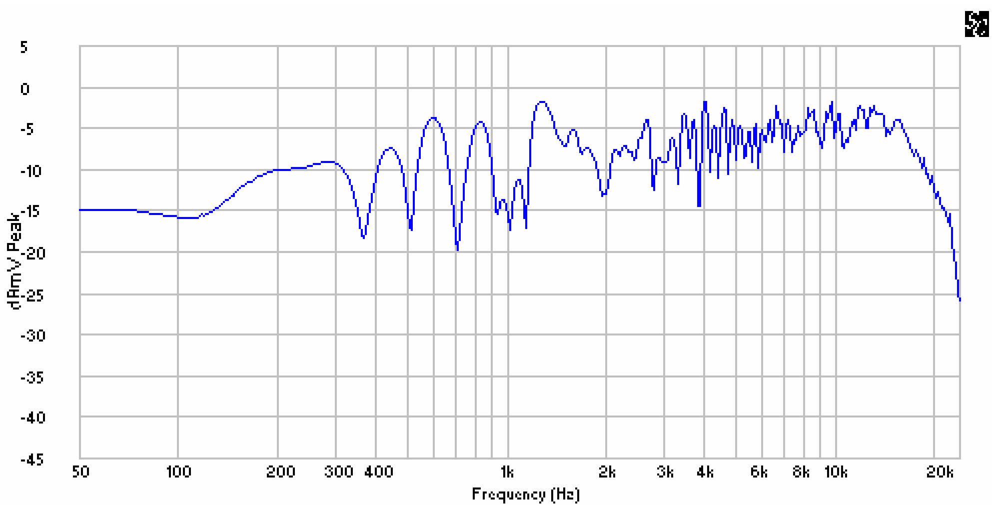

Fig. 5 Vehicle A - measured response for left-hand loudspeaker at desk operator position, first 9 ms.

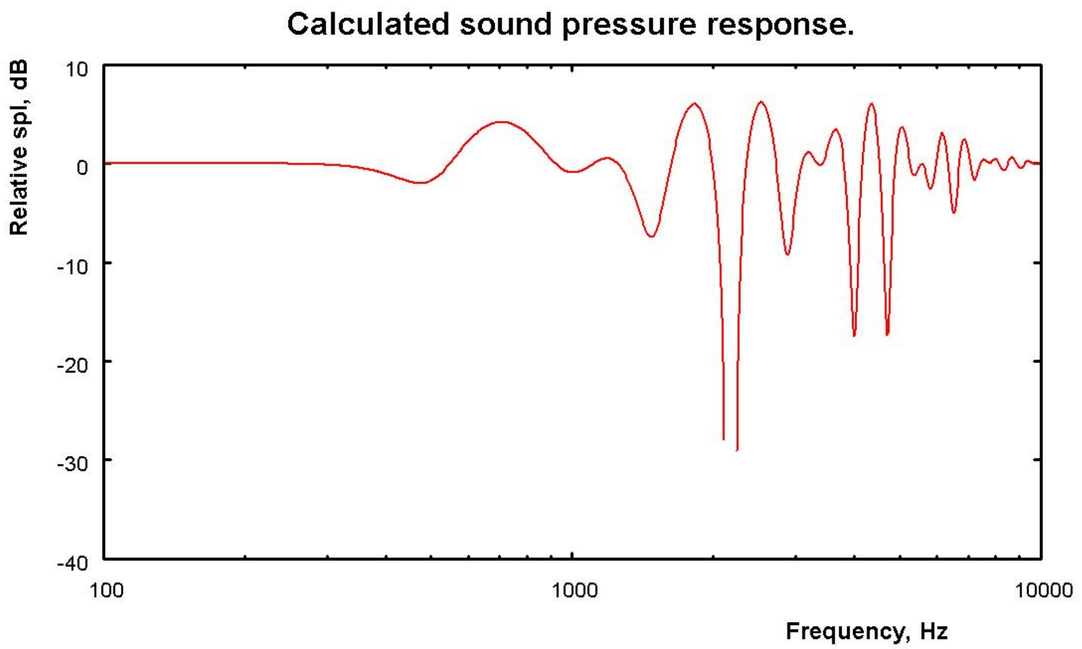

Fig. 6 Vehicle A - calculated response for first 1.5 ms, centre position.

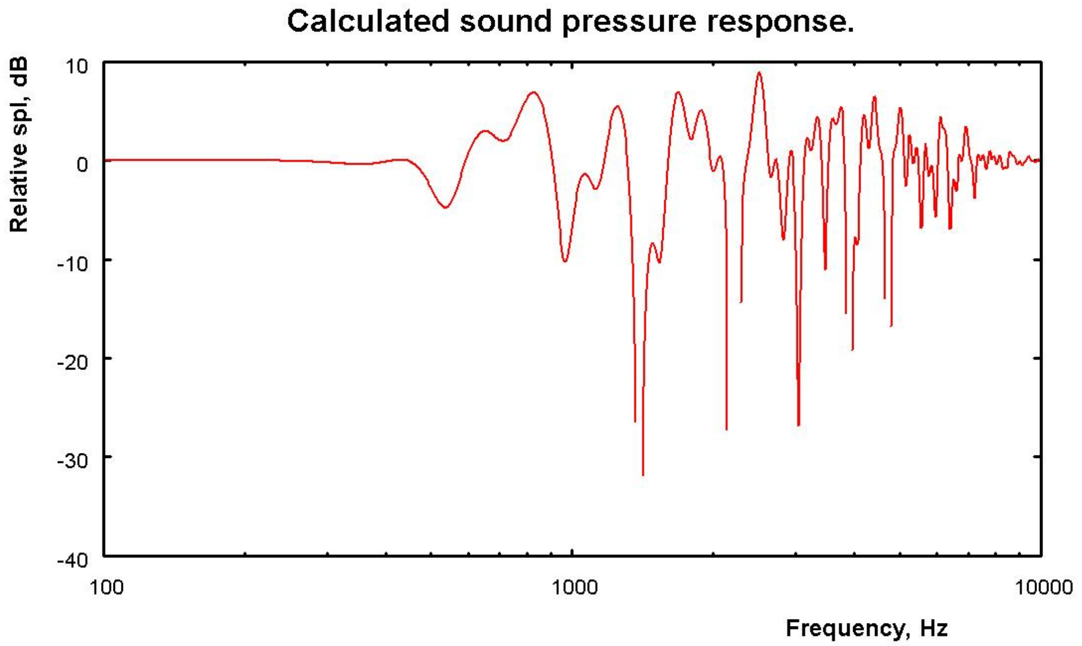

Fig. 7 Vehicle A - calculated response for first 9 ms., centre position.

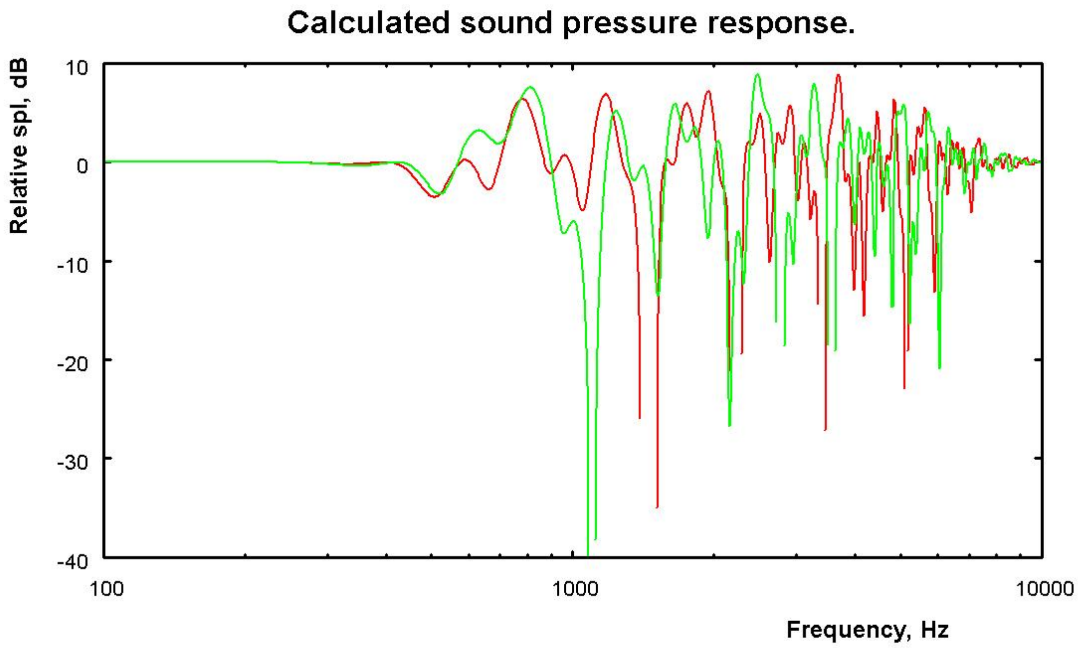

Fig. 8. Vehicle A - calculated responses for first 9 ms. 100 mm left and right of centre, equivalent to left and right ear positions for central listener.

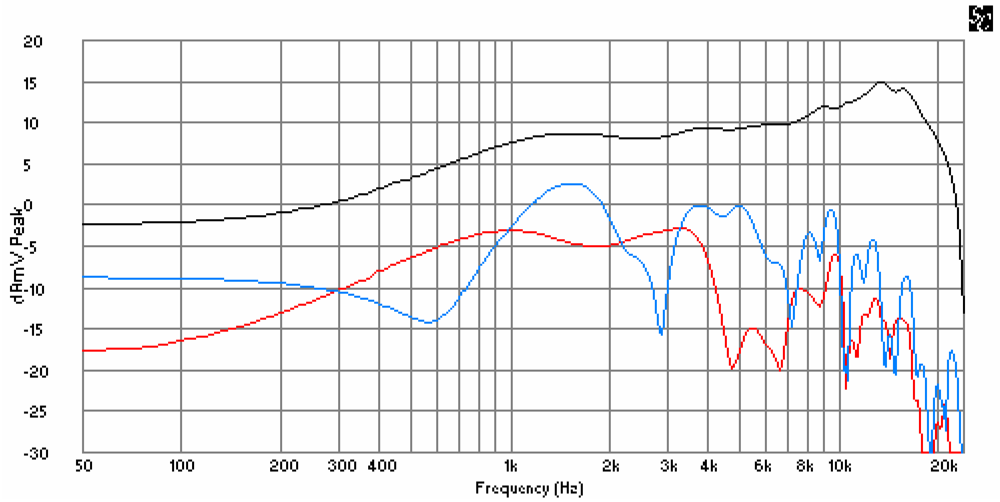

Fig. 9. Vehicle A - spectra for direct sound (black), opposite side wall (red) and ceiling (blue) reflections using main loudspeakers (1.5 ms time window).

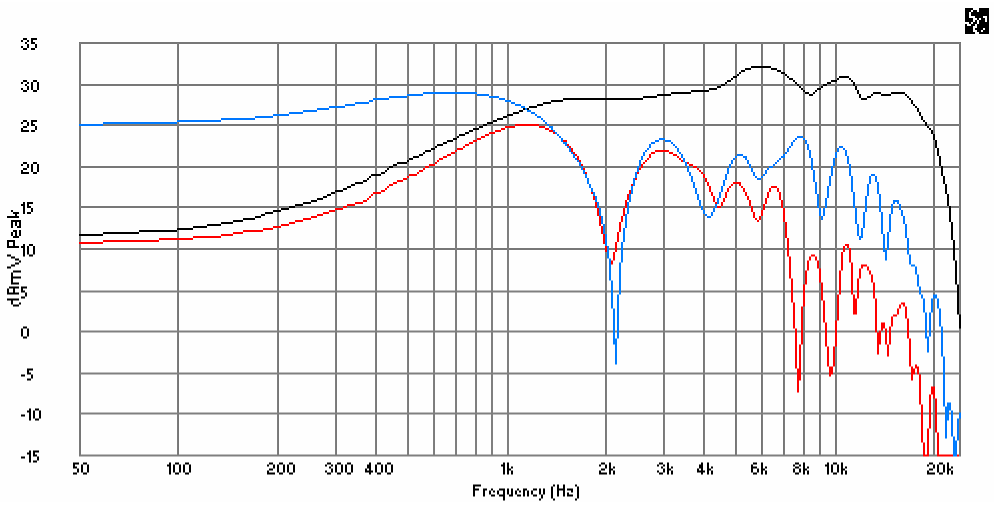

Fig. 10. Vehicle B - spectra for direct sound (black), opposite side wall (red) and ceiling (blue) reflections using main loudspeakers (2.0 ms time window).

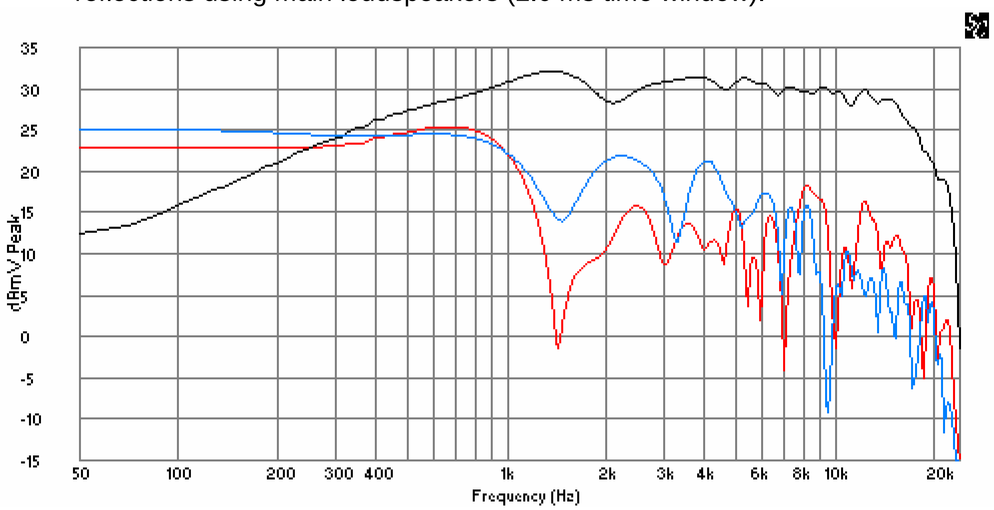

Fig. 11. Vehicle A - spectra for direct sound (black), opposite side wall (red) and ceiling (blue) reflections using near-field loudspeakers (3 ms time window).

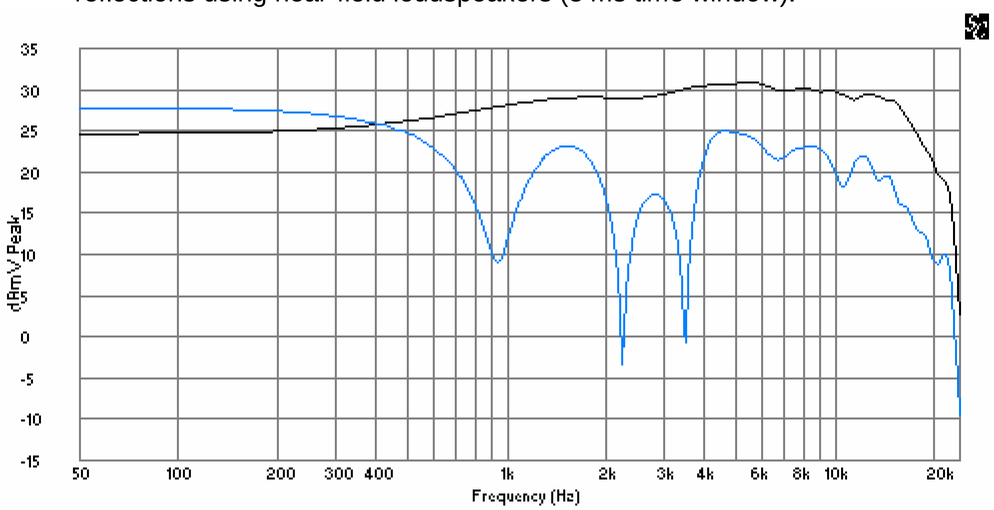

Fig. 12 Vehicle A - spectra for direct sound (black) and desktop (blue) reflections using near-field loudspeakers (1.5 ms time window).